

# **Document made available under the Patent Cooperation Treaty (PCT)**

International application number: PCT/EP05/050502

International filing date: 07 February 2005 (07.02.2005)

Document type: Certified copy of priority document

Document details: Country/Office: DE

Number: 10 2004 008 197.2

Filing date: 18 February 2004 (18.02.2004)

Date of receipt at the International Bureau: 15 March 2005 (15.03.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse



**Prioritätsbescheinigung über die Einreichung  
einer Patentanmeldung**

**Aktenzeichen:** 10 2004 008 197.2

**Anmeldetag:** 18. Februar 2004

**Anmelder/Inhaber:** Siemens Aktiengesellschaft, 80333 München/DE

**Bezeichnung:** Qualitätsorientierte Bewertung klinischer  
Studiendaten

**IPC:** G 06 F 19/00

**Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ur-  
sprünglichen Unterlagen dieser Patentanmeldung.**

München, den 19. Januar 2005  
**Deutsches Patent- und Markenamt**  
**Der Präsident**  
Im Auftrag

A handwritten signature in black ink, appearing to read "Hois".

Hois

Beschreibung

Qualitätsorientierte Bewertung klinischer Studiendaten

1. Die Qualität der Datenerhebung für klinische Studien,  
5 z.B. zur Zulassung neu entwickelter Medikamente, ist entscheidend für den Erfolg der Studie, und kann somit über Zulassung oder Nicht-Zulassung eines mit Milliardenaufwand entwickelten Medikamentes entscheiden.  
15 Datenerhebung umfasst in diesem Kontext das gesamte Spektrum medizinisch -klinischer Daten, von einer Patientenbefragung ("sind die Schmerzen stark, mittel oder schwach"), über eine Blutdruckmessung bis hin zu einer Röntgenaufnahme. Zwar werden im Rahmen der Studienprotokolle exakte Vorschriften erstellt, wie die Daten zu erfassen sind. Eine Kontrolle aller  
20 Datenerhebungsvorgänge, ob die Vorschriften eingehalten werden, ist jedoch nicht möglich. Auch lassen die Ausführungsvorschriften immer noch gewisse Spielräume offen. Letztendlich wird ein erfahrener Spezialist immer bessere Daten erheben als ein unerfahrener Neuling.  
25 Damit erhebt sich das Problem, wie die Qualität erhobener Daten für klinische Studien retrospektiv beurteilt werden kann. Weiterhin fehlen bisher Verfahren, wie mithilfe retrospektiver Bewertung der erhobenen Daten eine qualitätsabhängige Vergütung für die Datenerhebung bestimmt werden kann, oder wie für die künftige Auswahl der Partner zur Durchführung klinischer Studien (Investigatoren) eine Qualitäts-Benchmarking durchgeführt werden kann.
- 30 2. Eine standardisierte vergleichende Bewertung der Datenqualität erfolgt nicht, und es gibt auch noch keine Modelle zur erfolgsabhängigen Vergütung klinischer

Studien, bei denen die Vergütung von der Datenqualität abhängig gemacht wird.

3. Bei den meisten Daten, die im Rahmen klinischer Studien erhoben werden, ist die Qualität eines Einzeldatums nicht bewertbar. Werden jedoch Daten über große Patientenkollektive erhoben, so kann erwartet werden, dass gewisse statistische Größen zweier Datenmengen, die von unterschiedlichen, aber in ihrer Zusammensetzung vergleichbaren Patientenkollektiven erhoben werden, idealer weise gleich sein sollten.
- Als willkürliches Beispiel sei genannt: Wenn bei zwei Patientenkollektiven mit gleicher Alters-, Geschlechts- und ethnischer Verteilung der Blutdruck unter standardisierten vorgeschriebenen Bedingungen erhoben wird, muss erwartet werden, dass der Mittelwert des Blutdrucks über alle Patienten in beiden Gruppen gleich ist. Ist dies nicht der Fall, so muss ein systematischer Grund für den Unterschied vorhanden sein. Ein solcher Unterschied kann in einem fehlerhaft kalibrierten Messgerät, in einer unterschiedlich (bzw. falsch) durchgeführten Messverfahren (z.B. die Blutdruckmanschette wird zu weit unten am Arm angebracht), Messung nach körperlicher Anstrengung oder nach Ruhepause etc liegen.
- Ein solcher Unterschied kann aber auch in bisher nicht bedachten systematisch unterschiedlichen Zusammensetzung der Patientenkollektive liegen. Beispiel: die Kollektive kommen aus deutlich unterschiedlichen sozialen Verhältnissen. Das eine Einzugsgebiet ist New York's Bronx, das andere Beverly Hills in Los Angeles. Nun ist bekannt, dass Diabetes in sozial schwächeren Schichten häufiger ist, und dass die Grunderkrankung Diabetes zu Bluthochdruck führt, also mit großer Wahrscheinlichkeit zu unterschiedlichem mittlerem Blutdruck in zwei Patientenkollektivem mit deutlich unterschiedlichem

mittlerem sozialen Niveau. Dies mag ein Beispiel sein, das erfahrene Entwickler klinische Studienprotokolle vorab bei der Auswahl der Patientenkollektive hätten bedenken müssen. Es wird aber sicher zahlreiche weitere Kausalketten geben, die weniger offensichtlich sind, beim Studiendesign nicht bedacht wurden, und die Aussagekraft der erhobenen Daten deutlich beeinträchtigen.

Als weiteres Beispiel sei angeführt: wenn eine Messgröße unter gut standardisierten Bedingungen erhoben wird, sind meist glatte Messwertkurven zu erwarten, die mit einem geringen, aber unvermeidlichen Messwertrauschen behaftet sind. Beispiel: ein Patient bekommt ein Blutdrucksenkendes Mittel. Der Blutdruck wird täglich gemessen, und sollte im Verlauf von 4 Wochen aufgrund des neuen Medikaments tendenziell abnehmen (Fig. ). Bei unsachgemäßer Durchführung der Blutdruckmessung bzw. Abweichung von den standardisierten Bedingungen kann dieselbe Messkurve beim selben Patienten mit deutlich höherem Rauschen behaftet sein (Fig. ). Eine Methode der Qualitätsbewertung der Daten wäre also, die Streuung (das "Rauschen") jeder Messkurve bei allen Patienten zu berechnen, und für ein Patientenkollektiv den Mittelwert des Rauschens aller Messkurven zu bestimmen.

Das Gegenteil kann auch der Fall sein: jede Messkurve wird ein für das Messverfahren typisches Grundrauschen (in absoluter Höhe und auch im Rauschmuster, etwa Frequenzspektrum des Rauschens) aufweisen. Ist nun das mittlere Rauschen aus vielen Messkurven untypisch kleiner als ein erwarteter Mindestwert, so deutet dies ebenfalls auf einen systematischen Fehler bei der Messdatenerhebung hin, bis hin zu "geschummelten" Daten (Schätzwerte statt Messungen).

Es sind aus der Mathematik und Statistik zahlreiche Methoden bekannt, die für eine derartige Auswertung zum Vergleich zweier Datenmengen angewendet werden können. Ohne Einschränkung der Allgemeinheit seinen  
5 beispielhaft genannt: (geometrischer, arithmetischer) Mittelwert, Streuung, Variationskoeffizient, Parameter einer Gaußverteilung, Trendanalyse, Korrelationsanalyse etc. Auch Mustererkennungsmethoden können eingesetzt werden, wie zum Beispiel Methoden zur Erkennung und  
10 Charakterisierung von räumlichen Clustern in mehrdimensionalen Datensätzen.

Das vorgeschlagene Verfahren zur vergleichenden Qualitätskontrolle von zwei oder mehr Datenmengen aus  
15 zwei oder mehr Patientenkollektiven in klinischen Studien umfasst eine (beliebige) Kombination der folgenden Verfahrensschritte:

- Festlegen von Qualitätskontrollparametern für Messdaten, und ggf. von Erwartungswerten, Ober- und Untergrenzen für diese Parameter in den Studienprotokollen.
- Auswertung der Datenkollektive unter Berechnung der Qualitätskontrollparameter
- Vergleich der Qualitätskontrollparameter zwischen mehreren Datenkollektiven
- Festlegung der Zielvereinbarung für die Qualitätskontrollparametern mit den Investigatoren, und Regelung der Bezahlung für die klinische Studie in Abhängigkeit von den erreichten  
30 Qualitätskontrollparametern, entweder in absoluten Größen, oder als Prozentsatz vom besten Teilnehmer (Benchmarking)
- Anlegen von Datenbanken mit früher erhobenen Qualitätskontrollparametern zum Zwecke des Benchmarking:  
35 mit Hilfe der Datenbank wird ein Ranking der Qualität der Investigatoren durchgeführt: Diese Information kann

bei der Auswahl der Investigatoren für künftige Studien genutzt werden

- entlarven von betrügerisch gefälschter Datenerhebung (z.B. mittels technisch nicht möglichen Rauschspektren in den Daten)

- Ausschluss von Datenkollektiven aus der Studie, wenn diese nachweisbar vorbestimmten Qualitätskriterien, gemessen mit den Qualitätskontrollparametern, nicht genügen.

- Vorab-Analyse der Datenkollektive nach einem festgelegten Meilenstein (z.B. 10% der Patientenzahlen),

und Ausschluss von Investigatoren mit schlechter Performance, bzw. ggf Aufdecken von systematischen Fehlern in den Studienprotokollen bzw. der Zusammensetzung der Patientenkollektive, und Änderung des Studienprotokolls

- Erstellen eines Toolsets, das alle sinnvollen mathematisch-statistischen Methoden zur Ableitung von Qualitätsparametern enthält, und Anwenden des Toolsets retrospektiv auf 2 oder mehr Datenkollektive aus klinischen Studien; Aufdecken von systematischen Unterschieden in den Datenkollektiven. Verwendung der gefundenen systematischen Unterschiede entweder zur künftigen Verwendung als Qualitätskontrollparameter in Studienprotokollen, oder zur Aufdeckung von bisher nicht erkannten Kausalketten, die zu systematischen Unterschieden in Datenkollektiven führen. Des weiteren Ableitung von Grenzwerten für die Qualitätskontrollparameter in künftigen Studienprotokollen.

In einer sehr vorteilhaften Ausführungsform haben die Datenbanken der klinischen Daten ein standardisiertes Format, und das Toolset der mathematisch-statistischen Auswertemethoden dem Format angepasste standardisierte Schnittstellen, so dass zum Zwecke einer Auswertung eine

Benutzeroberfläche ausreicht, bei der der Benutzer nur mehr einen Algorithmus ("Streuung") und eine Datenkategorie ("Blutdruck") und eine Datenquelle (Studienname , Investigator) auswählen muss, und ohne weitere Eingaben, Formatierungen oder Datentransfers das Ergebnis der Auswertung erhält.

4. Die Qualität der Daten aus klinischen Studien wird mit Hilfe mathematisch-statistischer Methoden durch vergleichende Bewertung der abgeleiteten Qualitätsparameter aus zwei oder mehr Datenkollektiven bestimmt. Die gewonnene Information über die Datenqualität wird genutzt um

- ein Ranking der Investigatoren durchzuführen, und für künftige Studien die Investigatoren mit der besten Performance auszuwählen

- die Bezahlung für die Erhebung klinischer Daten von der Datenqualität abhängig zu machen

- die Qualität der Daten in einer Vorabanalyse zu quantifizieren, und die Studienprotokolle ggf. zu verbessern

- Datenkollektive oder Investigatoren, die den Qualitätskriterien nicht genügen, aus der Studie auszuschließen

Gemäß einem unabhängigen Anspruch werden Datenkollektive retrospektiv mit einem Toolset mathematisch-statistischer Methoden analysiert, um geeignete Qualitätskontrollparameter , deren typische Mittelwerte und Grenzwerte zu bestimmen, und um bisher unerkannte Kausalketten und Korrelationen für fehlerhafte Datenerhebung aufzudecken.

